# 2.0 TEST METHOD PROTOCOL COMPONENTS OF THE 3T3 AND NHK *IN VITRO* NRU TEST METHODS

The *Guidance Document* (ICCVAM 2001b) recommended that the following conditions be incorporated into any *in vitro* cytotoxicity protocol used to predict *in vivo* acute lethality:

- use a cell line (or primary cells) that divides rapidly
- use an initial seeding density that allows rapid growth throughout the exposure period
- apply reference substances only on cells in the exponential phase of growth
- use a reference substance exposure period at least the duration of one cell cycle
- use appropriate positive and vehicle control substances for which cytotoxicity, or lack of cytotoxicity, has been well characterized by the performing laboratory
- use solvents only at levels previously shown not to cause cytotoxicity to the cell system over the entire period of the assay
- use a well established measurement endpoint that has good interlaboratory reproducibility
- use tests compatible with 96-well plates and apparatus (i.e., spectrophotometers) that allow a quick and precise measurement of the endpoint
- use a progression factor in the concentration-response experiment that yields graded effects between no effect and total cytotoxicity

**Section 2.1** provides descriptions of the protocol applications to the NICEATM/ECVAM *In Vitro* Cytotoxicity Validation Study. **Section 2.2** provides details for performing the 3T3 and NHK NRU test methods and explains the rationale for various test method components. The basis for the selection of these *in vitro* cytotoxicity test methods is given in **Section 2.3** and proprietary aspects associated with this study are described in **Section 2.4**. **Section 2.5** discusses the basis for replicate and repeat tests. **Section 2.6** details the modifications and revisions made throughout all phases leading to the development of the final protocol used in Phase III of this validation study. **Section 2.7** shows the differences between the test methods used in this study and the test methods outlined in the *Guidance Document*.

**Sections 2.8** and **2.9** provide details on the solubility protocol for the reference substances used in to validate the two *in vitro* NRU cytotoxicity test methods.

These test method protocols were provided to the three cytotoxicity testing laboratories that participated in the NICEATM/ECVAM study (see **Section 5.6.3** for additional laboratory information):

- ECBC: The U.S. Army Edgewood Chemical Biological Center
- FAL: Fund for the Replacement of Animals in Medical Experiments (FRAME) Alternatives Laboratory
- IIVS: Institute for *In Vitro* Sciences

A fourth laboratory was used (BioReliance Corporation, Rockville, MD) to procure and distribute the coded reference substances and to perform solubility tests on all validation study reference substances prior to distribution to the cytotoxicity testing laboratories.

## 2.1 Overview of the 3T3 and NHK NRU Test Methods

The authors of the *Guidance Document* (ICCVAM 2001b) developed and presented a proposed 3T3 NRU protocol for use in a validation study based on the BALB/c 3T3 Cytotoxicity Test, INVITTOX Protocol No. 46 (available at the FRAME-sponsored INVITTOX database [http://embryo.ib.amwaw.edu.pl/invittox/]) which in turn was based on the Borenfreund and Puerner (1985) protocol, as elaborated on in Spielmann et al. (1991) and Spielmann et al. (1996).

The *Guidance Document* protocol also included revisions based on experience with a modification of another test, the 3T3 NRU Phototoxicity Test, INVITTOX Protocol No. 78, also available at the FRAME database. The Registry of Cytotoxicity (RC) regression for prediction of acute oral systemic rodent (rat and mouse) toxicity (Halle 1998; Spielmann et al. 1999) was included as the prediction model (see **Section 1.1.2**). The RC is a database of acute oral $LD_{50}$ values for rats and mice obtained from RTECS® and $IC_{50}$ values from *in vitro* cytotoxicity assays using multiple cell lines and cytotoxicity endpoints for chemicals with known molecular weights.

The NHK NRU test method protocol in the *Guidance Document* was based on a NRU test method by Borenfreund and Puerner (1984) using human epidermal keratinocytes (Heimann and Rice 1983) and was obtained from IIVS. Formulations for the media and solutions and general NHK cell culture techniques correspond to Clonetics® products from the CAMBREX Corporation. The authors of the *Guidance Document* expanded the IIVS protocol by adding details on equipment, media and reagent components, and experimental procedure.

The test method protocol components for the *in vitro* NRU cytotoxicity test methods used in the NICEATM/ECVAM study are very similar for both the 3T3 and the NHK cells (see **Figure 2-1**). The following procedures are common to both cell types:

- preparation of reference substances and positive control
- cell culture environmental conditions
- determination of test substance solubility
- 96-well plate configuration for testing samples
- range finder and definitive tests (48-hour exposure to the reference substance)
- microscopic evaluation of cell cultures for toxicity
- measurement of NRU
- data analysis

The main differences in the test methods are:

- the conditions of propagation of the cells in culture
- the cell growth medium components
- the application of reference substances to the 96-well plate (i.e., different volumes of reference substance solution)

The nature of the NRU response is described in **Section 1.3.1**. **Figure 2-1** provides an overview to the major steps for performance of the *in vitro* NRU cytotoxicity test methods.

**Figure 2-1 Major Steps for Performance of the NRU Test Methods in the *In Vitro* Cytotoxicity Validation Study**

(1) 3T3 cells or NHK cells are seeded into 96-well plates to form a sub-confluent monolayer (24 hours for 3T3 cells, 48-72 hours for NHK cells)

⇓

(2) Culture medium is removed (for 3T3 cells only)

⇓

(3) Reference substances in treatment medium are added to the cells; cells are exposed for 48 hours to the reference substance over a range of eight (8) concentrations

⇓

(4) Cells are evaluated microscopically for toxicity based on morphological alterations

⇓

(5) Treatment medium is removed; cells are washed once with Dulbecco's Phosphate Buffered Saline (D-PBS); Neutral Red (NR) dye medium is added (3T3 cells: 25 μg/mL NR dye; NHK cells: 33 μg/mL NR dye); plates are incubated for 3 hours

⇓

(6) NR medium is discarded; cells are washed once with D-PBS; NR desorbing fixative is added to the plates

⇓

(7) Plates are shaken for 20 minutes

⇓

(8) NR absorption is measured at optical density (OD) 540 ± 10 nm

⇓

(9) NRU is calculated as the % of control values to define $IC_{20}$, $IC_{50}$, and $IC_{80}$ reference substance concentrations (μg/mL)[1]

---

[1] $IC_{50}$ values are used for estimating the LD50 value of a reference substance. The $IC_{20}$ and $IC_{80}$ values were collected (as per request in the validation study's Statement of Work [SOW]) for possible use in estimating human lethal concentrations in blood.

### 2.1.1 The 3T3 NRU Test Method

*Initiating and Subculturing of 3T3 Cells*

(CCL-163, 3T3 BALB/c mouse fibroblast, clone 31, American Type Culture Collection [ATCC], Manassas, VA, USA)

Cryopreserved 3T3 cells are thawed, resuspended in a routine culture medium containing Dulbecco's Modification of Eagle's Medium (DMEM) supplemented with non heat-inactivated 10% newborn calf serum (NCS), transferred into tissue culture flasks (25 or 75 - 80 $cm^2$), and incubated at 37°C ± 1°C, 90% ± 5% humidity, and 5.0% ± 1% $CO_2$/air. When cells reach 50 – 80% confluency (as estimated from a visual inspection of cell density), they are removed from the flask by trypsinization. A single-cell suspension is added to new flasks for propagation and the cells are passaged/subcultured at least two times before seeding into 96-well plates for test assays. Subsequent passages may be maintained in culture for approximately two months (~18 passages) and used in NRU test methods. A new frozen ampule is thawed when needed and the above procedures are repeated. The protocols provide cell culture density guidelines for subculturing the cells and each laboratory determines the final seeding densities to achieve appropriate growth.

*Preparation of Cells for 96-well Plate Assays*

After achieving appropriate subculturing of cells, 100 µL of the cell suspension (2.0 – $3.0 \times 10^3$ cells/well) are placed in the appropriate wells and 100 µL of cell-free culture medium are dispensed into the peripheral wells (blanks). One plate per reference substance is prepared. The cells are incubated for 24 ± 2 hours and checked to be sure that approximately a half-confluent monolayer is attained at the time of reference substance application.

*Reference Substance Application*

After the appropriate incubation period, medium is removed and 50 µL of the routine culture medium with 10% NCS are added to each well. Then, 50 µL treatment medium containing the appropriate reference substance concentrations are added for a final concentration of 5% NCS. The cells are incubated for 48 ± 0.5 hours. At the end of the incubation period, the cells are microscopically evaluated for changes in morphology and their appearance is documented (as per Visual Observation Codes in the protocol) prior to measurement of the NRU of the cells.

### 2.1.2 The NHK NRU Test Method

*Initiating and Subculturing of NHK Cells*

(pooled primary neonatal foreskin cells, Clonetics® # CC-2507, lot # 1F0490N, CAMBREX Bio Science Walkersville, Inc., Walkersville, MD, USA)

Cryopreserved cells are thawed, resuspended in keratinocyte complete growth medium, transferred into tissue culture flasks (25 $cm^2$ without fibronectin-collagen coating), and incubated at 37°C ± 1°C, 90% ± 5% humidity, and 5.0% ± 1% $CO_2$/air. When cells reach 50 – 80% confluency (as estimated from a visual inspection of cell density), they are removed from the flask by trypsinization and prepared for subculturing into the 96-well plates. Keratinocytes are not subcultured beyond the second passage. Additional frozen ampule(s) are thawed as needed. The protocols provide cell culture density guidelines for establishing the cells out of cryopreservation and each laboratory determines the final seeding densities to achieve appropriate growth.

*Preparation of Cells for 96-well Plate Assays*

After appropriate subculturing of cells is achieved, 125 µL of the cell suspension (2.0 – $2.5 \times 10^3$ cells/well) are placed in the appropriate wells and 125 µL of cell-free culture medium are dispensed into the peripheral wells (blanks). One plate per reference substance is prepared. The cells are incubated for ~ 48 - 72 hours and checked to be sure that a monolayer of 20+% confluency (e.g., 20 – 50% confluency) is attained at the time of reference substance application.

*Reference Substance Application*

After the appropriate incubation period, 125 µL of the culture medium containing the appropriate reference substance concentrations are added to the test wells (the existing 125 µL of culture medium is not removed). The cells are incubated for 48 ± 0.5 hours. At the end of the incubation period, the cells are microscopically evaluated for changes in morphology and their appearance is documented (as per Visual Observation Codes in the protocol) prior to measurement of the NRU of the cells.

#### 2.1.3 Measurement of NRU for both 3T3 and NHK Test Methods

The treatment medium is removed from the 96-well plates, the cells are rinsed with phosphate buffered saline (PBS), 250 µL NR dye medium is added to the wells (25 µg NR/mL concentration for 3T3 cells, 33 µg NR/mL concentration for NHK cells), and the plates are incubated (37°C ± 1°C, 90% ± 5% humidity, and 5.0% ± 1% $CO_2$/air) for three hours. After incubation, the NR medium is removed, the cells are rinsed with PBS, and the desorb solution is applied. The plates are shaken on a microtiter plate shaker for 20 to 45 minutes to extract NR from the cells and form a homogeneous solution. The absorption (i.e., OD measurement) of the resulting colored solution is measured (within 60 minutes of adding the desorb solution) at 540 nm ± 10 nm in a spectrophotometric microtiter plate reader, using the blanks as reference. Data from the plate reader is transferred to a Microsoft® EXCEL® (Microsoft Corporation, Redmond, WA, USA) spreadsheet template (hereafter know as EXCEL® template) designed by the SMT and laboratories for statistical analyses for this study.

### 2.2 Descriptions and Rationales of the 3T3 and NHK NRU Test Methods

The protocols used in Phases I, II, and III of the validation study (**Appendices B and C**) are modifications of the protocols reported in the *Guidance Document* (ICCVAM 2001b, **Appendix D**). The SMT and the cytotoxicity laboratories provided comments and recommendations in the development of these protocols. The following information is specific to the NICEATM/ECVAM validation study.

#### 2.2.1 Materials, Equipment, and Supplies

*3T3 Cells*

3T3 cells (see **Section 2.1.1**), an immortalized mouse fibroblast cell line, were procured from the ATCC by IIVS at passage number 64. IIVS placed the cells in culture to expand the number of cells and cryogenically-preserved them as a pool at passage number 69. ECBC and FAL received frozen ampules of cells at passage number 69 from IIVS, propagated the cells, and cryopreserved multiple ampules of cells at a slightly higher passage number to establish a working cell bank (for each laboratory) for use throughout the study.

*NHK Cells*

These normal human epidermal keratinocytes are primary neonatal foreskin cells pooled from several donors and were obtained from CAMBREX Bio Science Walkersville, Inc. (**see Section 2.1.2**). IIVS reserved the specific lot of pooled cells (stored at CAMBREX) for use throughout the study by all laboratories. At each laboratory, cryopreserved NHK cells are thawed from a cryogenic ampule, seeded into culture flasks, propagated according to protocol, then trypsinized and seeded into 96-well plates. NHK cells are passaged only once (to the 96-well plates) and each new assay begins with fresh cells from the cryogenically preserved working bank if NHK cells in the culture flasks are too confluent according to protocol guidelines.

*Tissue Culture Materials and Supplies*

The 3T3 and NHK NRU test methods require general tissue culture materials and supplies (see **Appendices B-1 and B-2** [protocols] for formulations and concentrations of solutions and media). Both test methods use the same materials for solubility testing (**Section 2.8.1**). Freshney (2000) provides information on all aspects of cell culture including materials, supplies, and equipment needed. The following materials are needed for both test methods:

- trypsin (i.e., 0.05% trypsin)
- PBS
- Hanks' Balanced Salt Solution (HBSS) without $Ca^{2+}$ and $Mg^{2+}$
- NR dye
- glacial acetic acid
- dimethyl sulfoxide (DMSO)
- ethanol (ETOH)
- distilled water

*Culture Medium*

Medium for 3T3 cells consists of DMEM containing high glucose (4.5 gm/L) and supplemented with non heat-inactivated NCS, L-glutamine, penicillin, and streptomycin. The culture medium for NHK cells consists of Clonetics® keratinocyte basal medium (KBM®) supplemented with KBM® SingleQuots® (epidermal growth factor, insulin, hydrocortisone, antimicrobial agents, bovine pituitary extract) and Calcium SingleQuots® (calcium)[all from CAMBREX Corporation].

*Cell Culture Materials*

Laboratory items needed include the following:

- sterile, disposable tissue culture plasticware (e.g., 25 $cm^2$ - 75 $cm^2$ flasks, multiwell/microtiter plates [96-well], petri dishes) `
- cryogenic ampules
- pipettes, pipette tips
- multichannel solution reservoirs
- centrifuge tubes
- microporous sterilization filters
- general plastic containers
- glass tubes (for preparation of reference substance dilutions)

*Equipment*

Performance of the NRU test methods requires a laboratory equipped with a designated cell culture area. Essential equipment for cell culture work and the NRU test method includes:

- incubator (37ºC ± 1ºC, 90% ± 5% humidity, 5.0% ± 1% $CO_2$/air)
- laminar flow clean bench/cabinet (standard: "biological hazard")
- water bath (37ºC ± 1ºC)
- inverted phase contrast microscope
- centrifuge (capable of 220 x g)
- laboratory balance (capable of measuring to 10 mg)

- 96-well plate spectrophotometer (i.e., microtiter plate reader) equipped with 540 nm ± 10 nm filter
- shaker for microtiter plates
- cell counter or hemocytometer
- pipetting aid
- pipettes, pipettors (multi-channel and single channel, multichannel repeater pipette)
- waterbath sonicator
- refrigerator
- freezer
- cryostorage container (liquid nitrogen).
- magnetic stirrer
- antistatic bar ionizer
- personal computer
- osmometer
- pH meter

### 2.2.2 Reference Substance Concentrations/Dose Selection

Each laboratory prepares the reference substance immediately prior to testing (i.e., same day as test). Bulk solutions are not prepared for subsequent testing. The highest concentration of dissolved reference substance is identified using the solubility protocol and designated as the 2X stock solution. All reference substance dilutions for the assay are serially derived from the stock solution (see **Appendix D** [*Guidance Document*] for serial dilution methods).

*Range Finder Test*

A range finder test is the initial 3T3 and/or NHK NRU test method performed to determine starting doses for the main (definitive) test. The range finder test uses eight concentrations of the reference substance prepared by diluting the stock solution in log dilutions to cover a large concentration range. The highest concentrations applied to the cells are 10 mg/mL for reference substances dissolved in culture medium and 1 mg/mL in medium for reference substances dissolved in DMSO, unless precluded by the solubility of the reference substance.

ETOH was not used as a solvent in NRU test methods for any of the 72 reference substances in the NICEATM/ECVAM study.

If a range finder test does not generate enough cytotoxicity, then a second range finder test is conducted at higher doses, unless precluded by solubility. If solubility is an issue, then more stringent solubility procedures are employed to increase the stock concentration (to the maximum concentration specified in **Appendices B-1 and B-2**). If the test produces a biphasic response curve for NR uptake, then the doses selected for the subsequent definitive tests (concentration-response assays) cover the most toxic dose-response range that includes the range where 50% toxicity is first exceeded (see **Section 2.6.3** – *Unusual Dose-Response Curves*).

*Definitive Test*

In the following, because of its capacity to determine the $IC_{50}$ value of a test compound, the main test of the *3T3 and/or NHK* NRU test method will be referred to as the definitive test. The concentration closest to the calculated $IC_{50}$ value in the range finder test can serve as the midpoint of the eight concentrations tested in a definitive test. In the absence of other information (e.g., knowledge of slope for the toxicity curve), the recommended dilution factor is 1.47 ($\sqrt[6]{10}$), which divides a log into six equidistant steps (e.g., 10, 14.7, 21.5, 31.6, 46.4, 68.1, 100), as a starting dilution series. A progression factor of 1.21 ($\sqrt[12]{10}$) is regarded the smallest factor achievable and was the lowest dosing interval allowed in the validation study. The positive control chemical is tested similarly to the reference substances in the definitive test.

A successful definitive test is one that meets all of the test acceptance criteria as outlined in the protocol. Definitive tests were repeated as per the protocols if the test failed to meet all test criteria. **Section 2.5** addresses the basis for replicate testing.

If minimal or no cytotoxicity is measured in the dose range finding test, the maximum dose for a definitive test is as follows:

- Reference Substances Prepared in NHK or 3T3 Medium: the highest reference substance concentration applied to the cells in the definitive test is either 100 mg/mL (using 200 mg/mL 2X stock) or the maximum soluble dose. A review of the RC chemicals used in this study showed that, among water-soluble chemicals, glycerol had the highest reported $IC_{50}$ value (57 mg/mL). To capture this value during testing and that of other relatively non-toxic chemicals, the 100 mg/mL upper concentration limit was established.

- Reference Substances Prepared in DMSO: the highest test article concentration applied to the cells in the definitive test is either 2.5 mg/mL, or the maximum soluble dose.

### 2.2.3 NRU Endpoints Measured

*Neutral Red Uptake and Measurement*

After cells are exposed to the reference substance or the positive control chemical for the specified period, 3T3 or NHK cells are incubated with the NR dye for three hours, the dye is eluted from the lysosomes using a desorb solution, and the OD of the resulting colorimetric endpoint is measured using a spectrophotometric microtiter plate reader. The OD values are a reflection of the NRU by the cells. The greater the OD value is, the greater the NRU and the higher the percent viability[2] of the cells is in reference to the vehicle control (VC) wells. These OD data are transferred to the EXCEL® template. The mean OD values of the six replicate values (six wells [minimum of four] in the 96-well plate) per test concentration are used to determine relative cell viability by calculating its percentage of the mean NRU of all VC values on the same plate.

*Determination of $IC_{50}$, $IC_{20}$, and $IC_{80}$ Values*

The $IC_{50}$ values are determined from the concentration response using a Hill function which is a four parameter logistic mathematical model relating the concentration of the reference substance to the response (typically following a sigmoidal shape). Information on

[2] Vehicle control wells are considered to have 100% cell viability (i.e., all cells are alive). Cell viability in other test wells is referenced to the vehicle control value.

modifications to the Hill function used in later phases of the validation study may be found in **Section 2.6.3**.

Data from the EXCEL® template were transferred to a template designed by the SMT for a commercially available statistical software program (GraphPad PRISM® 3.0, GraphPad Software, Inc., San Diego, CA, USA – hereafter known as PRISM® template) to generate the inhibitory concentrations $IC_{50}$, $IC_{20}$, and $IC_{80}$ reported as μg/mL of reference substance in solution. $IC_{20}$ and $IC_{80}$ data were collected for potential use in designing a prediction model for estimating human lethal blood concentrations.

### 2.2.4 Duration of Reference Substance Exposure

The SMT and laboratory representatives reevaluated the reference substance exposure duration recommended in the *Guidance Document* (ICCVAM 2001b) before initiating the NICEATM/ECVAM study. The *Guidance Document* recommends an exposure of 24 hours for the 3T3 cells and 48 hours for the NHK cells. The results from a cytotoxicity study by Riddell et al. (1986) show large differences in cytotoxicity in 3T3 cells induced by some chemicals depending on whether an exposure duration of 24 or 72 hours was used. IIVS conducted studies to evaluate the effect of exposure duration (24, 48, and 72 hours) on the sensitivity of 3T3 cells to six chemicals selected from the list in Riddell 1986. Since the closest fit to the RC regression line (Halle 2003) occurred when 48-hour exposure duration was used, this exposure duration is used in the standardized protocol for 3T3 cells (see **Appendix E**). In addition, IIVS evaluated the sensitivity of NHK cells to the same six chemical using exposure durations of 48 and 72 hours. To make a comparison with the RC regression, the 11 chemicals recommended by the *Guidance Document* were tested in both cell types using the same exposure durations. IIVS scientists concluded that the optimum exposure duration for both cell types was 48 hours (Curren et al. 2003). The SMT concurred and revised the exposure duration in the 3T3 protocol to 48 hours.

### 2.2.5 Known Limits of Use

*Solubility/Volatility*

*In vitro* cytotoxicity test methods are inadequate for substances that cannot be dissolved in media, DMSO, or ETOH at a sufficiently high concentration to induce cytotoxicity in excess of 50%. Some reference substance dilutions in this study had precipitates in various 2X concentrations prior to dilution for application to the test plates. Precipitates were observed in a number of test plates after addition of solutions to the cultures and at the end of testing (1X solutions [see **Section 3.5** and **Table 5-11**]). Volatility was detected for a number of reference substances during the range finder tests by observance of cross contamination of wells (i.e., high cytotoxicity in some VC wells). Some volatility was controlled by using plate sealers during the definitive tests (see **Section 2.6.3** – *Testing Volatile Reference Substances*). Plate sealers could be used during the range finder tests if the laboratory suspected that the reference substance might be volatile. However, use of plate sealers requires additional laboratory skills and highly volatile reference substances are difficult to test even with the use of plate sealers. Additionally, some test substances (e.g., organic solvents) may react chemically with the plastic plate sealers. Also, chemicals that are unstable or exothermic in water cannot be adequately tested with these test methods.

*Biokinetic Determinations*

The Workshop report (ICCVAM 2001a) provides discussions on the role of the kinetics of a chemical *in vivo* vis a vis its acute systemic toxicity.

> "Results obtained from *in vitro* studies in general are often not directly applicable to the *in vivo* situation. One of the most obvious differences between the situation *in vitro* and *in vivo* is the absence of processes regarding absorption, distribution, metabolism and excretion (i.e., biokinetics) that govern the exposure of the target tissue in the intact organism. The concentrations to which *in vitro* systems are exposed may not correspond to the actual situation at the target tissue after *in vivo* exposure. In addition, the occurrence of metabolic activation and/or saturation of specific metabolic pathways or absorption and elimination mechanisms may also become relevant for the toxicity of a compound *in vivo*. This may lead to misinterpretation of *in vitro* data if such information is not taken into account.

> Therefore, predictive studies on biological activity of compounds require the integration of data on the mechanisms of action with data on biokinetic behavior."

Biokinetic determinations were not specifically addressed in this study.

*Organ-Specific Toxicity*

The Workshop report also addresses concerns about which *in vitro* test methods can adequately predict organ-specific toxicity and identifies the organ systems in which failure after acute exposure could lead to lethality (liver, central nervous system, kidney, heart, lung, and hematopoietic system). Each system is reviewed individually and a five-step *in vitro* testing scheme (as opposed to a single *in vitro* test method) that could act as a test battery that may eventually be used as a replacement for *in vivo* acute toxicity testing is proposed.

- Step 1 of the proposed *in vitro* scheme recommends performing a physico-chemical characterization and biokinetic modeling.
- Step 2 promotes the use of a basal cytotoxicity test method (e.g., 3T3 and NHK NRU test methods).
- Step 3 calls for a test to determine the potential that metabolism will mediate the basal cytotoxicity effect.
- Step 4 is to assess the test substance's effect on energy metabolism.
- Step 5 is to assess the ability of the substance to disrupt epithelial cell barrier function (ICCVAM 2001a).

Organ-specific toxicity and metabolic effects were not tested in this study.

### 2.2.6 Nature of Response Assessed

Neutral red is a weakly cationic, water-soluble dye that stains living cells by readily diffusing through the plasma membrane and concentrating in lysosomes. The intensity of the dye in culture is directly proportional to the number of living cells. In addition, since altering the cell surface or the lysosomal membrane by a toxicological agent causes lysosomal fragility and other adverse changes that gradually become irreversible, cell death and/or inhibition of cell growth decreases the amount of neutral red taken up by the culture (see **Section 1.3.1**).

### 2.2.7 Appropriate Vehicle, Positive, and Negative Controls

*Positive Control (PC)*

The *Guidance Document* recommended sodium lauryl sulfate (SLS, Chemical Abstracts Service Reference Number [CASRN] 151-21-3) as an appropriate PC chemical for *in vitro* cytotoxicity test methods (ICCVAM 2001b). SLS is frequently used for this purpose and historical data are available (e.g., Spielmann et al. 1991). A PC test plate was included with each run of any 3T3 and/or NHK NRU test method assay and was treated the same as any reference substance assay plate.

The acceptable range for the PC $IC_{50}$ was based on the statistical approach recommended in the *Guidance Document*. Initially, in Phase Ia of the validation study, the 3T3 and NHK tests were considered acceptable if the $IC_{50}$ was within the 95% confidence interval of an historical mean $IC_{50}$ value. The SMT decided that the test acceptance criterion for the $IC_{50}$ for Phase III of the validation study (for both cell types) was 2.5 standard deviations of the mean SLS $IC_{50}$ data obtained during Phases I and II. The exception to this was the FAL NHK data, where only the Phase II data were used as the basis for establishing the acceptable range for the PC. SLS data produced at FAL during Phase I was not used due to a protocol change in culturing the cells (see **Section 2.6.2** – *Resultant protocol changes for Phase II*). The historical mean, standard deviation, and acceptance limits were determined separately for each laboratory (see **Table 5-2**).

*Vehicle Control (VC)*

For the NICEATM/ECVAM validation study, the VC consisted of complete DMEM (see **Appendix B-1**) for 3T3 cells and complete Clonetics® KBM® (see **Appendix B-2**) for NHK cells for reference substances dissolved in medium. For reference substances dissolved in DMSO, the VC consisted of medium with the same amount of solvent as that used in the reference substance concentrations that are applied to the 96-well test plate (i.e., 0.5 % [v/v]).

*Negative Control*

A negative control was not incorporated into the NRU test methods. The SMT and study directors decided that the vehicle control would be used in place of a negative control.

### 2.2.8 Acceptable Ranges of Control Responses

The *Guidance Document* established the use of the absolute value of the $OD_{540}$ value of NRU obtained in the untreated VC to indicate whether the cells seeded in the 96-well plate have grown exponentially with a normal doubling time during the assay. A mean $OD_{540} \geq 0.3$ was recommended as the acceptable range of VC responses and was made a test acceptance criterion for both cell types. Protocols for Phases II and III provide a range of OD values for use as guidance in each phase of the study.

**Table 2-1 Vehicle Control $OD_{540}$ Ranges**

| Phase | $OD_{540}$ Range - 3T3 | $OD_{540}$ Range - NHK | Notes |
|---|---|---|---|
| Ia | $\geq 0.3$ and $\leq 1.1$ | $\geq 0.3$ and $\leq 1.1$ | Test Acceptance Criterion |
| Ib | $\geq 0.30$ and $\leq 0.80$ | $\geq 0.60$ and $\leq 1.70$ | Test Acceptance Criterion |
| II | $\geq 0.103$ and $\leq 0.813$ | $\geq 0.35$ and $\leq 1.50$ | Target Range (not criterion) |
| III | $\geq 0.103$ and $\leq 0.813$ | $\geq 0.205$ and $\leq 1.645$ | Target Range (not criterion) |

In Phase III, 99.5% (914/919) of all 3T3 mean VC OD values and 97% (913/944) of all NHK mean VC OD values were within the target range. Most OD values out of the ranges were from range finding tests and were usually the result of volatile reference substances affecting the VC cells nearest the highest reference substance concentration.

*VCs as Quality Control*

To check for systematic cell seeding errors and potential volatility issues, untreated VCs were placed both at the left side (row 2) and the right side (row 11) of the 96-well plate (see **Appendices B-1 and B-2**). Volatile reference substances generally affect the left side VC (closest to the highest reference substance concentration). The test acceptance criterion was that the left and the right mean of the VCs did not differ by more than 15% from the mean of all VCs. This criterion was used in all phases of the study for reference substances and PC test plates.

### 2.2.9 Nature of Experimental Data Collected

Each laboratory maintained a Study Workbook to document all aspects of this study. All raw data from cell culture procedures (e.g., cell growth, application of reference substances, NRU test method, etc.) and all solubility studies were recorded in the workbook.

*NRU OD Measurements*

At the conclusion of the NRU desorb step, the OD of the resulting colored solution in each well of the 96-well plates was measured at 540 ± 10 nm in a spectrophotometric microtiter plate reader. Raw OD data from the plate reader was transferred to the EXCEL® template. The template converts the raw data (six wells/reference substance concentration) to derived data by subtracting the mean blank value (two wells/reference substance concentration) associated with each reference substance concentration. The VCs had a total of 12 test wells and 20 blanks. The corrected OD values were referenced to the mean VC OD value and a relative viability (% of VC) was determined for each test well. The percent viability values was then transferred to the PRISM® template for calculation of the $IC_{20}$, $IC_{50}$, and $IC_{80}$ values.

*Type of Data Collected*

Originals of the raw data (the Study Workbook and computer printouts of absorbance readings from the plate reader) and copies of other raw data such as instrument logs were collected and archived under the direction of the Study Director according to Good Laboratory Practice (GLP)-compliant procedures.

The Study Director/technicians entered the following information to the EXCEL® template:

- raw data: OD values from microtiter plate reader
- testing identification for: test facility, chemical code, study number, 96-well plate number, experiment number
- reference substance preparation: solvent used, solvent concentration in dosing solutions, highest stock concentration, dilution factor, pH of 2X dosing solutions, medium clarity/color, presence/absence of precipitate in 2X solutions, PC concentration range

- cell line/type: cell supplier, lot number, cryopreserved passage number, passage number in assay
- cell culture conditions: medium/supplements and supplier and lot numbers, serum concentrations
- test acceptance criteria: acceptable number of values on each side of the $IC_{50}$ (i.e., number of points > 0 and ≤ 50% viability and > 50 and < 100% viability), acceptable % difference for the VCs, acceptable Hill function $R^2$ value (coefficient of determination) for the PC, and calculated $IC_{50}$ concentration for the PC
- timeline: dates for cell seeding, dose application, $OD_{540}$ determination
- test results: mean corrected $OD_{540}$ value, Hill function $R^2$ value, logs of $IC_{20}$, $IC_{50}$, and $IC_{80}$ (PRISM® template presents data as logs of the $IC_x$; EXCEL® converts values to $IC_x$ in μg/mL)
- visual observations: protocol codes for cell culture conditions for all reference substance concentrations (i.e., relative level of cell cytotoxicity, cell morphology, presence of precipitate)

2.2.10 Type of Media for Data Storage

Raw data from the NRU cytotoxicity test methods was saved in the EXCEL® template file format provided by the SMT for further analysis of the concentration-response (percent viability calculations). The derived test method data were stored electronically. All EXCEL® and PRISM® files were copied and transferred to compact disks. NICEATM and the laboratories printed copies of all data sheets (stored at NICEATM and at the testing facilities). Copies were also included in the final reports.

2.2.11 Measures of Variability

Each 96-well plate used in the NRU test methods has three main measures of variability.

1) Each plate contains VCs on each end of the plate (columns 2 and 11). The percent difference between each column and the mean of both columns is calculated and was used as a test acceptance criterion. If the difference was

greater than 15%, then the test was rejected by the Study Director. This value is an indicator of reference substance volatility and potential cell seeding errors.

2) A mean relative viability was determined for each concentration along with the standard deviation and % coefficient of variation (CV).
3) Macros were included in the EXCEL® template to perform an outlier test (Dixon and Massey 1981) on data in each well of the test plate. Extreme values at the 99% level were highlighted and could be removed to improve curve fit. The decision as to whether or not to remove outliers was made by the Study Director.

Other test-to-test measures of variability were considered for this study.

- Each set of assays include a PC plate. If the SLS PC data did not meet test acceptance criteria, then all tests associated with that PC were rejected. The SMT recommended testing a manageable number of definitive test plates (e.g., 4 – 6) with each PC to avoid rejection of reference substance NRU assays that are unacceptable due only to a PC failure. In this validation study, 4.2% of all definitive tests performed were rejected only because the PC failed (i.e., the PC $IC_{50}$ was outside the acceptable confidence limits).
- Standard deviations and CVs were determined for mean $IC_{50}$ values from replicate testing of the same substance. Replicate testing included three definitive tests per reference substance, each performed on a different day.

2.2.12 <u>Methods for Analyzing NRU Data</u>

A calculation of cell viability expressed as NRU was made for each concentration of the reference substance by using the mean NRU of the six replicate values (minimum of four acceptable replicates wells) per test concentration. This value was compared with the mean NRU of all VC values (provided VC values have met the VC acceptance criteria). Relative cell viability was expressed as percent of untreated VC. Raw OD data from the microtiter plate reader was transferred to the EXCEL® template for performance of these calculations. Where possible, the eight concentrations selected for each reference substance tested ranged from no effect up to 100% toxicity.

The $IC_{20}$, $IC_{50}$, and $IC_{80}$ values were determined from the concentration-response by using the PRISM® template and applying a Hill function to the data. The $IC_{20}$ and $IC_{80}$ values were calculated for use in the development of a human prediction model resulting from this study.

### 2.2.13 Decision Criteria for Classification of Reference Substances

The 3T3 and NHK NRU test methods were not used to classify reference substances in hazard categories but rather to aid in setting the starting dose for acute systemic toxicity assays (i.e., the Up and Down Procedure [UDP], the Acute Toxic Class method [ATC], the Fixed Dose Procedure [FDP]). The RC regression formula (i.e., the prediction model) was used to predict an $LD_{50}$ value from an NRU $IC_{50}$ value. The RC compilation (Halle 2003) contains *in vitro* cytotoxicity information on 347 chemicals (i.e., one average $IC_{50x}$ value/chemical based on multiple reports in the literature) with corresponding *in vivo* acute oral $LD_{50}$ values (mmol/kg) for rats (282 values) or mice (65 values) from RTECS (See Halle 2003 for the RC data). **Section 6** addresses the potential of using the *in vitro* NRU cytotoxicity test methods for predicting the GHS hazard category.

### 2.2.14 Information and Data Included in the Test Report

*Test and Control Substances*

(Laboratories in this study worked only with coded reference substances and could not know the specific reference substance information.)

- chemical name(s) such as the structural name used by the CASRN, followed by other names, if known
- the CASRN, if known
- formula weight, if known
- purity and composition of the substance or preparation (in percentage(s) by weight)
- physicochemical properties (e.g., physical state, volatility, pH, stability, chemical class, water solubility)
- treatment of the test/control substances (solubility efforts) prior to testing, if applicable (e.g., vortexing, sonication, warming, grinding)

- stability, if known

*Information Concerning the Sponsor and the Test Facility*

- name and address of the sponsor, test facilities, study director, and laboratory technicians
- justification of the test method and protocol used

*Test Method Integrity*

- the procedure used to ensure the integrity (i.e., accuracy and reliability) of the test method over time (e.g., use of the PC data)

*Criteria for an Acceptable Test*

- acceptable VC differences (between each column and the mean of both columns)
- acceptable concurrent PC ranges based on historical data
- number of cytotoxicity points on either side of the $IC_{50}$ (i.e., number of points $> 0$ and $\leq 50\%$ viability and $> 50$ and $< 100\%$ viability)

*Test Conditions*

- experimental start and completion dates
- details of test procedure used
- test concentration(s) used
- cell type used
- description of any modifications of the test procedure
- reference to historical data of the model (e.g., solvent and positive controls)
- description of evaluation criteria used

*Results*

- tabulation of data from individual test samples (e.g., $IC_{50}$ values for the reference substance and the PC, reported in tabular form, including data from replicate repeat experiments as appropriate, and means and the standard deviation for each experiment)

*Description of Other Effects Observed*

- for example, cell morphology, precipitate, NR crystals

*Discussion of the Results*

*Conclusion*

*Quality Assurance (QA) Statement for GLP-Compliant Studies*

- This statement indicates all inspections made during the study, and the dates any results were reported to the Study Director. This statement also serves to confirm that the final report reflects the raw data.

During this study, testing at IIVS and ECBC, the GLP-compliant laboratories, followed additional reporting requirements provided in the relevant guidelines (e.g., OECD 1998; EPA 2003a, 2003b; FDA 2003).

Standard forms for data collection, EXCEL® and PRISM® templates, were developed by the SMT and laboratories. The solubility test form was derived from a standard form provided by IIVS. The EXCEL® template was an adaptation of a template format presented in the *Guidance Document*.

## 2.3 Basis for Selection of the *In Vitro* NRU Cytotoxicity Test Methods

As stated in **Section 1**, Workshop 2000 participants recommended that the approach proposed by ZEBET (Halle 1998; Spielmann et al. 1999) be used for rapid adoption so that data could be generated to establish its usefulness with a large number of chemicals (ICCVAM 2001a). To assist in the adoption and implementation of the ZEBET approach, several workshop participants wrote the *Guidance Document* (ICCVAM 2001b). NICEATM and ECVAM used this document as the basis of test method protocol development and designed the validation study to evaluate the performance of the 3T3 and NHK NRU test methods.

### 2.3.1 *Guidance Document* Rationale for Selection of *In Vitro* NRU Cytotoxicity Test Methods

The *Guidance Document* (ICCVAM 2001b) provides basic protocols for using *in vitro* NRU basal cytotoxicity test methods as the means to predict a starting dose for *in vivo* acute lethality assays. The protocols take advantage of the relationship between *in vitro* $IC_{50x}$ values and *in vivo* $LD_{50}$ values derived from the RC for 347 chemicals (Halle and Spielmann

1992; Halle 2003). The 3T3 NRU and NHK NRU test method protocols used in the NICEATM/ECVAM validation study were derived from the document. Guidance was also provided for qualifying these tests for use with the RC regression to predict the starting dose.

The 3T3 NRU test method has been used most frequently in formal validation programs, all of which were aimed at evaluation of cytotoxicity in predicting eye irritancy. Large-scale studies include Phases I, II, and III of the Cosmetic, Toiletry, and Fragrance Association (CTFA) validation program (Gettings et al. 1991, 1992, 1994a, 1994b); the German eye irritation validation study (Spielmann et al. 1991, 1993, 1996); the European Commission/British Home Office (EC/HO) eye irritation validation study (Balls et al. 1995); and the European Cosmetic Toiletry and Perfumery Association (COLIPA) eye irritation study (Brantom et al. 1997). The 3T3 NRU Phototoxicity Test, a modification of the 3T3 NRU test, has been fully validated (Spielmann et al. 1998a,b), and has gained regulatory acceptance. See **Section 9** for comparison of these studies to this validation study.

### 2.3.2 *Guidance Document* Rationale for Selection of Cell Types

The Workshop (ICCVAM 2001a) concluded that there are no significant differences between the basal cytotoxicity results obtained using permanent mammalian cell lines, primary human cells, or using the $IC_{50x}$ approach of Halle and Spielmann (Halle 2003; Spielmann et al. 1999; Halle and Spielmann 1992). The Workshop recommended that near-term *in vitro* studies designed to reduce and refine animal testing in acute lethality tests should follow the ZEBET approach of using basal cytotoxicity assays in conjunction with the RC database. This can be one of the factors used to identify appropriate starting doses for *in vivo* acute lethality studies, as described by Spielmann et al. (1999).

*Cell Types for Basal Cytotoxicity Testing*

Established rodent (rat and mouse) cell lines were recommended because:

- it was assumed that such cells would give the best prediction of rodent (rat and mouse) acute lethality

- the use of an immortalized standard cell line that is easy to grow and readily available for *in vitro* cytotoxicity testing would hasten the generation of a database that can be used to analyze the usefulness of this approach

Human cells also offer potential advantages. An analysis of the RC rodent (rat and mouse) acute lethality data relative to cytotoxicity data generated using human cell lines in the MEIC program showed that both human and rodent cells were highly correlative ($R^2$=0.90) (ICCVAM 2001). A long-term advantage of using human cells is that the human cell cytotoxicity data derived from *in vitro* cytotoxicity testing can be added to human toxicity databases to facilitate the development of test methods that may later better predict acute human lethality.

*Differentiated Cells for Metabolic Capabilities*

The *Guidance Document* explained why highly differentiated cells were not used in the basal cytotoxicity assays. Such cells may not give the best prediction of acute lethality for the large variety of chemicals likely to be tested for acute toxicity (Ekwall et al., 1998). For example, to eliminate the possibility of metabolic activation or inactivation of chemicals, neither hepatocyte nor hepatoma cytotoxicity data were included in the RC database. This does not preclude the use of hepatocytes in future studies, however, either to estimate cytotoxicity or to investigate the effect of metabolism or cell-specific toxicity (Seibert et al., 1996). Hepatocytes are essential to investigations of metabolism-mediated toxicity (Seibert et al., 1996).

The Workshop participants agreed that the current *in vitro* basal cytotoxicity tests do not take into account metabolism-mediated toxicity. Simple predictive systems (*in vitro* or *in silico*) must be developed for early identification of those substances likely to be metabolized to more toxic or less toxic species than the parent chemical (e.g., Fentem et al., 1993; Seibert et al., 1996; Curren et al., 1998; Ekwall et al., 1999). Participants concluded that the available *in vitro* assays require further development to accurately predict acute lethality (i.e., $LD_{50}$). See **Section 3.3.4 – Metabolism** for metabolic information on the NICEATM/ECVAM reference substances.

*Historical Testing*

Historical data exists for 3T3 cells including data from controlled and blinded validation studies (Gettings et al. 1991, 1992, 1994a, 1994b; Spielmann et al. 1991, 1993, 1996; Balls et al. 1995; Brantom et al. 1997). Human NHK or fibroblasts have also been used in validation studies for basal cytotoxicity test methods with good results (Willshaw et al. 1994; Sina et al. 1995; Gettings et al. 1996; Harbell et al. 1997). See **Sections 5**, **6**, **7**, **8**, and **9** for data generated for the NICEATM/ECVAM validation study.

### 2.4 Proprietary Components of the *In Vitro* NRU Cytotoxicity Test Methods

The only proprietary components used in these test methods are the NHK cells and the NHK basal culture medium obtained from CAMBREX Clonetics®. All other components are readily available through various scientific product suppliers. The NHK cells consisted of pooled donor primary neo-natal foreskin keratinocytes from an unidentified source. The use of this specific supplier ensured that the laboratories would have access to the same source of keratinocytes throughout the entire validation study. Keratinocytes from other sources are acceptable if they meet the growth requirements identified in the protocols.

The contents of the NHK basal culture medium are proprietary, but the formulation is based on a commercially available basal medium (MCDB 153 formulation). This medium was chosen since it was recommended by the laboratories for use with the CAMBREX Clonetics® NHK cells and would be available for the laboratories throughout the study. Other media are acceptable for the NRU test methods if they meet the performance standards prescribed in the media prequalification protocol and achieve parity with the CAMBREX Clonetics® products (see **Appendix B-4** and **Section 2.6.3** – *Inadequate Cell Growth in NHK Medium*).

### 2.5 Basis for Number of Replicate and Repeat Experiments for the 3T3 and NHK NRU Test Methods

The NICEATM/ECVAM study protocols required each laboratory to test the reference substances in at least one range finding test using a log dilution factor and in at least three definitive tests on three different days using a smaller dilution factor than used in the range finding test. Assays were performed over a number of days to assess day-to-day variability.

Laboratories tested each coded reference substance until three definitive tests met the test acceptance criteria. Additional testing was often dictated by:

- chemical issues (low toxicity, volatility, insolubility, and precipitation)
- PC failure
- technical difficulties such as NR crystal formation

A stopping rule for insoluble reference substances was incorporated into the protocols to prevent infinite retesting:

> "If the most rigorous solubility procedures have been performed and the assay cannot achieve adequate toxicity to meet the test acceptance criteria after three definitive tests, then the Study Director may end all testing for that particular chemical."

## 2.6 Basis for Modifications to the 3T3 and NHK NRU Test Method Protocols

### 2.6.1 Phase Ia: Laboratory Evaluation Phase

All protocol revisions were implemented during Phase Ia unless otherwise stated.

*NR Dye Crystals*

NR dye crystals formed in the 96-well test plates in both NRU test methods when used at 50 μg/mL (OD values measured in the blanks increased from ~ 0.05 to 0.10). Troubleshooting efforts explored incubating the NR medium overnight, centrifuging, filtering, and reducing the concentration of NR dye. The laboratories performed tests using a reduced NR concentration of 33 μg/mL. Since there were no differences in results between tests with 50 μg/mL and tests with 33 μg/mL NR, the SMT accepted tests with both concentrations.

*Protocol Revision:* The NR dye concentration was reduced to 33 μg/mL for both cell types.

*3T3 Cell Growth*

Cell growth for 3T3 cells was slower than expected in that the cells required more time in culture after seeding cells from the cryogenically-preserved pool into culture vessels to obtain the proper density.

*Protocol Revision:* 3T3 cells must be passaged 2-3 times after thawing before reference substance application/toxicity evaluation. The protocol also emphasized attainment of the percent cell confluency required for both cell types prior to reference substance application rather than the amount of time in culture.

*NHK Cell Growth*

The NHK cells also had an additional growth problem that manifested as a ring of dead/dying cells around the center of the wells. Troubleshooting efforts included evaluating various brands of 96-well plates and eliminating the change of medium prior to reference substance treatment. All laboratories participated in evaluating the effect of changing (i.e., refeeding) or not changing (i.e., no refeeding) the medium by performing a small study with SLS, the PC. Tests were performed 1) after refeeding the cells with fresh medium, and 2) by adding SLS to the medium already on the cells. Control ODs were generally higher in the tests in which the medium was not replenished, but SLS sensitivity was unchanged (see **Table 2-2**). The SMT accepted both tests with refeeding and those without refeeding as long as they met the test acceptance criteria.

*Protocol Revision:* Step 2 of the NHK NRU test method was eliminated (change of medium prior to addition of reference substance). The volume of medium with cells placed into the 96-well plates was changed from 250 μL/well to 125 μL/well.

**TABLE 2-2 REFEEDING/NO REFEEDING DATA**

| | ECBC[1] | | IIVS[2] | | FAL[3] | |
|---|---|---|---|---|---|---|
| | **Refeed** | **No Refeed** | **Refeed** | **No Refeed** | **Refeed** | **No Refeed** |
| Number of Test Plates | 4 | 4 | 6 | 6 | 2 | 4 |
| Mean Abs. OD (VC) | 0.265 | 0.621 | 0.885 | 1.12 | 1.41 | 1.24 |
| Standard Deviation (SD) | 0.151 | 0.322 | 0.057 | 0.033 | 0.127 | 0.430 |
| SLS $IC_{50}$ (µg/mL) | 3.33 | 3.23 | 3.41 | 3.49 | 6.21 | 8.14 |
| SLS $IC_{50}$ SD | 0.47 | 0.61 | 0.58 | 0.39 | 0.88 | 0.40 |

[1]*Edgewood Chemical Biological Center*
[2]Institute for In Vitro Sciences
[3]FRAME Alternatives Laboratory

The FAL laboratory could not get satisfactory levels of NHK cell adherence to the 80-cm$^2$ culture flasks when seeded with thawed cells (one ampule) from the cryogenically-preserved pool of cells.

*Protocol Revision (FAL only):* Culture flasks were to be coated with fibronectin-collagen to promote adherence.

*OD Limits*

VC control OD limits (OD value must be ≥ 0.3 and ≤ 1.1 as related in the protocols) were frequently unattainable in both test methods. Study Directors reported that the cells were adequately responsive and were neither senescent nor 100% confluent. The SMT withdrew the VC control OD limits as a test acceptance criterion.

*Protocol Revision for Phase Ib:* OD data from all laboratories, a review of cell responsiveness (i.e., dose response data), and the ability of each test to pass the other acceptance criteria were analyzed for both cell types and new OD ranges were calculated as guidelines for each cell type.

*Precipitate Formation*

During solubility testing, precipitates were occasionally observed in the 3T3 medium but not in the NHK medium at the same reference substance concentrations. Some liquid reference substances (e.g., 2-propanol) caused precipitation in the 3T3 medium only. The precipitates were attributed to the serum in the 3T3 medium rather than insoluble reference substance.

*Protocol Revision:* The reference substance was dissolved in 3T3 medium without NCS. Then, for reference substance exposure, the dissolved 2X reference substance was added to medium containing 10% NCS to reach the final 5% NCS and 1X reference substance concentrations.

*Dilution Factor*

Once a range finder test had been performed, the definitive test assays were to be performed using a $\sqrt[6]{10} = 1.47$ dilution scheme centered on the $IC_{50}$. The laboratories sometimes deviated from the protocols and used dilution factors other than the required one. The SMT accepted data generated using dilution factors other than the recommended 1.47 for definitive tests if all other test acceptance criteria were met. The use of smaller dilution factors generally increased the number of points between 10 - 90% viability and the precision of the $IC_{50}$ calculation was improved.

*Protocol Revision:* The $\sqrt[6]{10} = 1.47$ dilution scheme was presented as a suggestion and was not a criterion for test acceptance after Phase Ia.

*Test Acceptance Criteria*

*The test acceptance criteria for Phase Ia were:*

- the $IC_{50}$ for SLS was within the 95% CI of the historical PC mean established by the Test Facility (*not applicable to Phase Ia*)
- mean OD values of the left and right VCs (columns 2 and 11 in the 96-well test plate) did not differ by more than 15% from the mean of all VC OD values
- at least two calculated cytotoxicity values, one on either side of the $IC_{50}$, between 10 and 90% viability (*added after commencement of Phase Ia*)
- Hill function coefficient of determination $R^2 > 0.9$ or $0.8 < R^2 < 0.9$ and curve fit was evaluated on a case by case basis for acceptability by the SMT (*added after commencement of Phase Ia*); (note: this determination would be made by the Study Director in non-validation studies)

- $OD_{540}$ of VCs (with blank subtracted) was ≥ 0.3 and ≤ 1.1 (*rescinded after commencement of Phase Ia*)

### 2.6.2 Phase Ib: Laboratory Evaluation Phase

*NR Crystal Formation*

FAL and ECBC routinely observed NR crystals forming in the 96-well test plates in 3T3 assays at 33 µg/mL NR. All laboratories tested 25 and 33 µg/mL NR concentrations and 2- and 3-hour exposure durations to determine which exposure duration would provide optimal NRU without crystal formation. In addition to determining whether NRU had reached a plateau at these concentrations and durations, the laboratories also tested SLS to determine whether sensitivity to SLS differed under these conditions. Crystals were observed only at 33 µg/mL NR when present for 3 hours. **Figure 2-2** shows that the average OD results were very similar for the concentrations and durations tested. **Figure 2-3** shows that the SLS $IC_{50}$ was approximately the same at these concentrations and durations. To minimize changes for the Phase III protocol, the SMT and laboratories agreed to use 25 µg/mL NR for three hours in the subsequent protocol revisions for the 3T3 test method. The NR concentration for the NHK NRU test method remained at 33 µg/mL.

*Protocol Revision for Phase II*: The NR concentration for the 3T3 NRU test methods was changed to 25 µg/mL NR for the three-hour incubation. Revised methods for preparation of the NR dye solution included filtration of the solution, maintenance of the solution at 37ºC, and application of the NR dye solution to the cells within 15 minutes after removing from 37ºC. Cells should be observed during the NR incubation period of the 3T3 and/or NHK NRU test method assays to monitor possible crystal formation.

**Figure 2-2 Optical Density with NR Concentration and Duration**

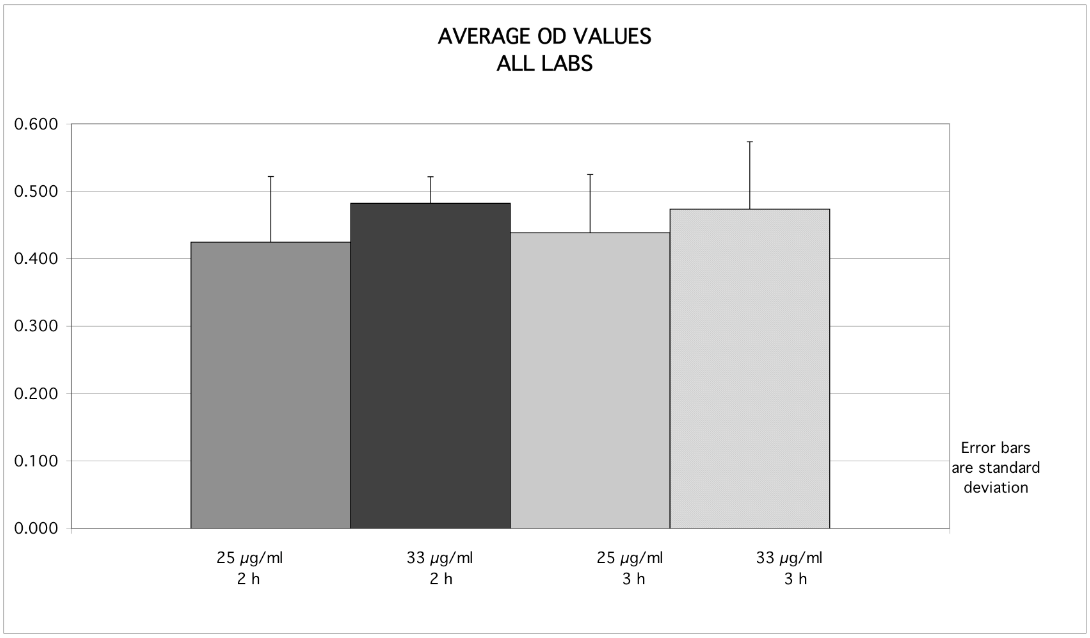
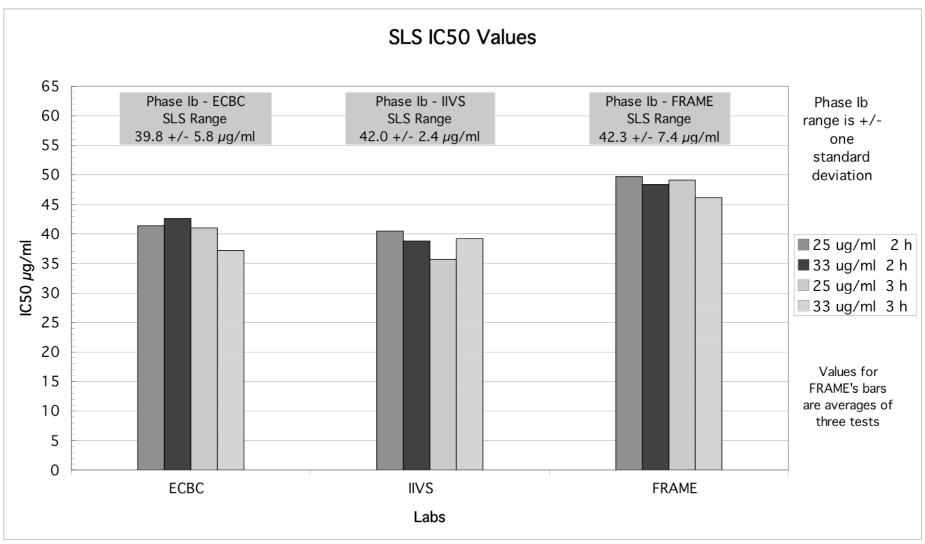



**Figure 2-3 SLS $IC_{50}$ for Each NR Concentration and Duration**




*Heating Reference Substance Solutions*

The laboratories had difficulty with the solubility of arsenic trioxide. Mechanical applications for solubilizing reference substances into culture medium were reviewed and revised.

*Protocol Revision for Phase II*: The range for duration of heating the reference substance solution was increased from 5 – 10 minutes to 5 – 60 minutes.

*OD Readings*

OD readings were frequently lower than acceptance criteria for the VC wells.

*Protocol Revision for Phase II:* The OD range was eliminated as a test acceptance criterion. The OD data from the VCs in the laboratories for both cell types was used to calculate OD ranges to serve as guidelines (see **Section 2.2.9**).

To adjust for potential reference substance interference with NR dye, the reference substance was added to the blank wells that were used to generate the background OD at 540 nm that was subtracted from the reference substance concentration ODs. Each reference substance concentration was applied to six wells containing cells and to two blank wells without cells.

*Laboratory Error Rates*

The SMT suggested that FAL needed additional guidance to become more GLP-like (e.g., improve documentation) and to improve performance (i.e., fewer test failures and errors) throughout Phases Ib and II. The SMT compiled a list of the errors (e.g., transcriptional and typographical errors in the data sheets) and error rates (number of tests with errors/number of tests) for the existing Phase Ib data and provided the information to each laboratory (see **Table 2-3**). IIVS management sponsored a weeklong laboratory training exercise at the IIVS facilities so that FAL technicians would have exposure to a GLP laboratory environment. ECBC was invited to participate and all three testing laboratories shared information and thereby harmonized procedures during the training exercise. Harmonization of the laboratory procedures illustrated the need to make additional protocol revisions.

**Table 2-3 Error Rates[a] in Phase Ib by Laboratory and Test Method**

| Laboratory | NRU Test Method | |
|---|---|---|
| | 3T3 | NHK |
| ECBC | 1/9 (10%) | 4/17 (23%) |
| FAL | 42/45 (93%) | 12/29 (41%) |
| IIVS | 1/20 (5%) | 1/20 (5%) |

[a]Number of tests with errors/total number of tests (some data files had more than one error)

*Resultant protocol changes for Phase II*

*The protocol changes include:*

- use multi-channel repeater pipettes for plating cells in the 96-well plates, dispensing plate rinse solutions, NR medium, and desorb solution, but not for dispensing reference substances to the cells; repeater pipettes are not accurate enough to deliver equal quantities of the reference substance solution to the wells
- use 8-channel reservoirs for applying dosing solutions to the wells so multi-channel single delivery pipettes could be used
- use a standardized length of time that HBSS rinses remain on the cell monolayers in flasks during the cell subculturing step
- protect plates from high light levels during the shaking step for NR extraction; all laboratories will cover plates (e.g., with aluminum foil) during this step
- allow plates to stand for at least five minutes after the shaking step is complete and break any bubbles observed in the wells before measuring OD
- change the seeding density range for 3T3 NRU test method from $2.5 \times 10^3$ cells/well to $2 - 3 \times 10^3$ cells/well
- change NHK culture flask size (at FAL) from 80-cm$^2$ (for start-up of cryopreserved cells) to 25-cm$^2$ (same as other laboratories) and discontinue using a fibronectin-collagen coating

*Test Acceptance Criteria*

*Criteria were modified as follows:*

- the $IC_{50}$ for SLS (PC) is within 2 SDs (approximately 95%) of the historical mean established by each laboratory in Phase Ia (originally used the 95% confidence interval)
- mean OD values of the left and right VCs (columns 2 and 11 in the 96-well test plate) do not differ by more than 15% from the mean of all VC OD values
- at least one calculated cytotoxicity value is between 10 and 50% viability and one calculated cytotoxicity value between 50 and 90% viability
- Hill function $R^2 > 0.9$ <u>or</u> $0.8 < R^2 < 0.9$ and curve fit is evaluated on a case by case basis for acceptability by the SMT (note: this determination would be made by the Study Director in non-validation studies)
- VC OD criteria are based on Phase Ia data (mean ± two SDs): 0.3-0.8 for the 3T3 test method, and 0.6-1.7 for the NHK NRU test method (*rescinded after commencement of Phase Ib*)

### 2.6.3 <u>Phase II: Laboratory Qualification Phase</u>

All revisions were implemented <u>during</u> Phase II unless otherwise stated.

*Testing Volatile Reference Substances*

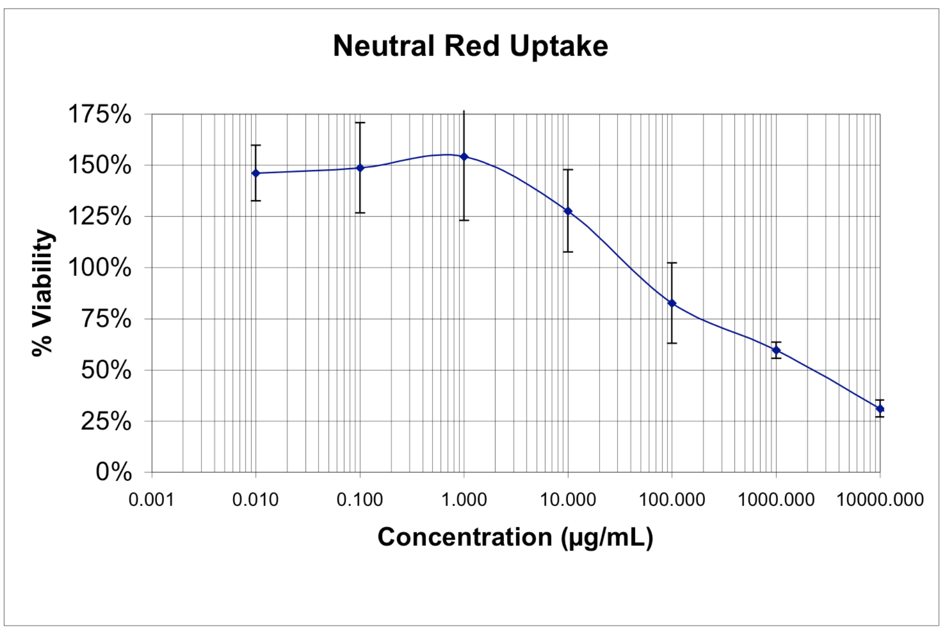
When 2-propanol was tested according to the protocol, vapors from the highest concentration wells contaminated the adjacent VC and appeared to affect some lower concentration wells (i.e., the wells exhibited unexpectedly reduced levels of NRU). An example dose-response curve is shown in **Figure 2-4**. The tests for which such contamination was present failed the VC criterion. When lower concentrations were used to avoid contaminating the VC adjacent to the highest concentration, toxicity was inadequate to produce an $IC_{50}$. To address this problem, IIVS repeated their tests using film plate sealers, which isolated all wells from each other, and obtained acceptable results. Based on these data, the SMT recommended the use of film plate sealers to the other laboratories to test 2-propanol.

FAL had previous experience using mineral oil as a cell culture cover to keep volatile reference substances from escaping and provided 2-propanol test data where mineral oil had been added to each well. The FAL showed that the average oil vs. film $IC_{50}$ values were not significantly different. However, there was less variability in the film sealer data than the mineral oil data so the SMT decided on the use of plate sealers.

A general indicator of volatility issues in the NRU test methods is the percent difference in the mean OD values for the two VC columns on the test plate. If the difference is greater than 15%, then reference substance volatility is suspected, especially if the VC adjacent to the highest test concentration had a significantly reduced OD value. Volatility may be an issue for reference substances with a specific gravity of less than 1. **Table 5-11** lists the study reference substances that had volatility issues in the NRU test methods.

> *Protocol Revision*: The SMT included the use of film sealers to test suspected volatile compounds in the Phase III protocols.

**Figure 2-4 Representative Dose-Response for 2-propanol in a 3T3 Range Finder Test**



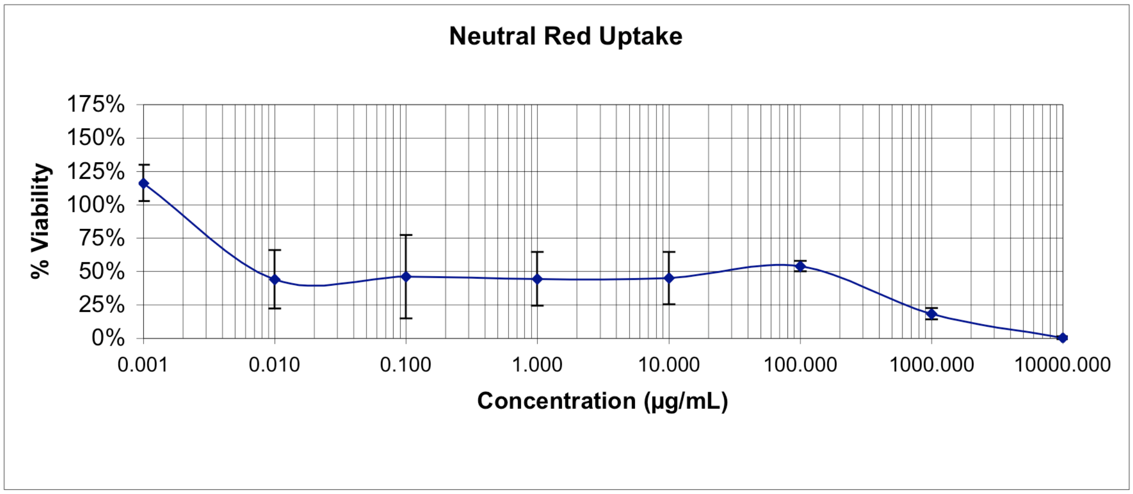

%Difference of the two VC columns from the average VC was 63%. Mean corrected OD for VC1, adjacent to the highest 2-propanol concentration was 0.070, while that for VC2, adjacent to the lowest 2-propanol concentration, was 0.310. The 100% viability of the mean VCs shifted the toxicity curve such that lower concentrations of 2-propanol seem to have viability percentages much greater than the VCs.

*Unusual Dose-Response Curves*

Some laboratories observed unusual dose-response curves for aminopterin and colchicine. When the range finder tests produced a biphasic response (see **Figure 2-5** for an example), the SMT advised the laboratories to focus the definitive tests on the lowest concentrations that produced responses around 50% viability. In the definitive tests, they noted that no matter how much reference substance was used, viability was not reduced to 0% (see **Figure 2-6**). This effect with colchicine was very reproducible across laboratories in the NHK NRU test method, but only FAL achieved this type of response with colchicine in the 3T3 NRU test method. Aminopterin produced a similar dose response in the NHK NRU test method at ECBC and FAL, but not at IIVS. In the 3T3 NRU test method, only FAL obtained an unusual response with aminopterin.

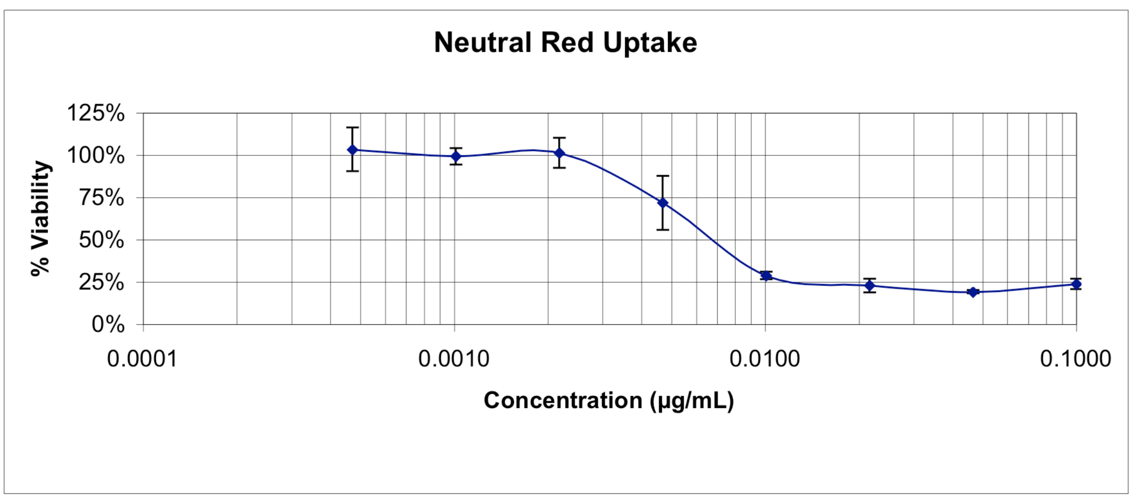
The SMT assumed the unusual dose-responses with these reference substances were due to their mechanisms of action. Colchicine binds to microtubular protein and interferes with function of mitotic spindles, which arrests cell division (NLM 2003). Aminopterin blocks the use of folic acid by the cells, which kills cells during the S phase of the cell cycle by inhibiting metabolism, RNA production, and protein synthesis (NLM 2002). The variability of results among the laboratories may be due to cells in the culture populations being in different cell cycle phases when reference substance was applied to the cultures. Application of reference substance to the cell systems is based on the cells being at a certain monolayer confluency that assures the cells are in exponential growth phase. A subjective visual observation of the cell cultures determines time point 0 for the reference substance exposure period for the NRU test method.

**Figure 2-5 Representative Dose-Response for Aminopterin in a NHK Range Finder Test**




Representative dose-response for aminopterin in a NHK range finder test. Laboratories were instructed to focus definitive tests (concentration-response assays) on the lowest doses that produced 50% viability.

**Figure 2-6 Representative Dose-Response for Aminopterin in a NHK Definitive Test**




Representative dose-response for aminopterin in a NHK definitive test (concentration-response assay). %Viability did not reach 0%.

*Hill Function*

The Hill function used in the previous phases of this study was defined as follows:

$$Y = \text{Bottom} + \frac{\text{Top} - \text{Bottom}}{1 + 10^{(\text{logIC50} - X)\text{HillSlope}}}$$

where Y= response, X is the logarithm of dose (or concentration), Bottom is the minimum response, Top is the maximum response, $\log IC_{50}$ is logarithm of X at the response midway between Top and Bottom, and HillSlope describes the steepness of the curve.

Since the unusual dose-responses did not fit the Hill function well, $R^2$ values often failed the acceptance criterion. To obtain a better model fit, the Bottom parameter was estimated without constraints (the previous practice was to use Bottom = 0). However, when Bottom ≠ 0, the $EC_{50}$ reported by the Hill function was not the same as the $IC_{50}$ since the Hill function relies on $EC_{50}$ defined as the point midway between Top and Bottom. Thus, the Hill function calculation using the Prism® software was rearranged to calculate the concentration corresponding to the $IC_{50}$ as follows:

$$X = \log EC_{50} - \frac{\log\left(\frac{Top - Bottom}{Y - Bottom} - 1\right)}{HillSlope}$$

where X is the logarithm of concentration at 50% response, $\log EC_{50}$ is logarithm of concentration at the response midway between Top and Bottom, Top is the maximum response, Bottom is the minimum response, Y = 50 (i.e., 50% response), and HillSlope describes the steepness of the curve.

IIVS performed the recalculations for their colchicine tests in the NHK NRU test method, but the SMT performed the necessary recalculations for the other laboratories. Tests that were recalculated by the SMT are noted in the data summaries.

*Protocol Revision*: The protocol was revised to state that if a range finding test produces a biphasic curve, then the concentrations selected for the subsequent tests should cover the most toxic dose-response range.

*Insoluble Reference Substances*

Lithium carbonate was insoluble in 3T3 medium. Only ECBC was able to expose 3T3 cells to sufficient lithium carbonate to produce three tests that passed the test acceptance criteria. Precipitate was reported for two of those tests in the wells at the three highest concentrations. Since the third highest concentration, 510.2 μg/mL, was approximately the $IC_{50}$ (average was 564 μg/mL), the true $IC_{50}$ for lithium carbonate may actually be lower than that calculated and therefore the $LD_{50}$ value would be underestimated. The data were not discarded.

*Protocol Revision for Phase III*: The protocol was revised to allow an increase in the solubility stirring/rocking duration in an incubator from 1 to 3 hours if cytotoxicity in the range finder test was limited by solubility. Also, a **Stopping Rule for Insoluble Chemicals** was added (see **Section 2.5**)

*Inadequate Cell Growth in NHK Medium*

IIVS and FAL had several NHK NRU test method assay failures that were attributed to poor cell growth. FAL found that medium/supplement lot combinations that performed poorly in the NHK NRU test method performed well for the laboratory's research on corneal cell cultures. The SMT compiled information from the laboratories on the KBM® and SingleQuot® lot numbers that the laboratories were using along with their assessment of NHK cell growth. The information was distributed to the laboratories to identify the lots that produced adequate growth. The SMT also obtained quality assurance and quality control test results from CAMBREX Clonetics® on the lots of KBM®, but the information provided was inadequate for determining how the medium would perform in the NHK NRU test method.

*Resolution*: A protocol for prequalifying the medium was developed (see **Appendix B-4**). For Phase III, the SMT asked IIVS to prequalify new lots of KBM® and SingleQuots® for use by all laboratories.

*Performance Standards for Media to Support NHK Growth*

A prequalification-of-medium protocol (**Appendix B-4**) was developed and IIVS performed several tests of different lots of medium and supplements to find various combinations that maintained the typical growth characteristics of cells in this study. The laboratories then reserved samples of these acceptable lots at CAMBREX so that the supply of media would not be interrupted due to unavailability of the materials.

*Test Acceptance Criteria for Prequalifying Media*

- $R^2$ (coefficient of determination) value calculated for the Hill model fit (i.e., from PRISM® software) is ≥ 0.85
- Difference between the mean of all VCs and (a) the left mean VC, and (b) the right mean VC is ≤ 15%
- At least one point > 0 % and ≤ 50.0 % viability and at least one point > 50.0 % and < 100 % viability
- After meeting all other acceptability criteria, the SLS $IC_{50}$ must be within the historical range established by the laboratory (i.e., mean SLS $IC_{50}$ ± 2.5 standard deviations)

*Other Criteria for Prequalifying Media (for consideration by a Study Director)*

- General culture observations: rate of proliferation; percent confluence; number of mitotic figures per field; colony formation; distribution of cells; absence or presence of contamination
- Cell morphology observations should include overall appearance (e.g., good, fair, poor), and presence of abnormal cells
- Mean corrected $OD_{540-550}$ of the VCs
- Cell morphology and confluence of the VCs at the end of the 48-hour treatment.
- Cell doubling time (determined by the laboratory for first time use of the NRU test method [prior to testing with SLS])

*Test Acceptance Criteria for Phase II*

- $IC_{50}$ for SLS (PC) is within 2.5 SDs of the historical mean established by the Test Facility (*Phases Ia and Ib*)
- Mean OD values of the left and right VCs (columns 2 and 11 in the 96-well test plate) do not differ by more than 15.0 % from the mean of all VC OD values (*change in decimal point only*)
- At least one calculated cytotoxicity value ≥ 10.0 % and ≤ 50.0 % viability and at least one calculated cytotoxicity value ≥ 50.1 % and ≤ 90.0 % viability (*change in decimal point only*)
- $R^2 \geq 0.90$. Test fails if $R^2 < 0.80$. If the $R^2 \geq 0.80$ and $< 0.90$, the SMT evaluates the model fit (note: this determination is made by the Study Director in non-validation studies)

### 2.6.4 Phase III: Laboratory Testing Phase

The changes below were made in the Phase III protocols as a result of the experience in Phase II.

*Cytotoxicity Values Around the $IC_{50}$*

Obtaining at least one calculated cytotoxicity value > 0 % and ≤ 50.0 % viability and at least one calculated cytotoxicity value > 50.0 % and < 100 % viability may be difficult or unattainable for reference substances with a steep dose response.

*Protocol Revision*: The test acceptance criterion was qualified so that tests with only one point between 0 and 100 % were acceptable if the smallest practical dilution factor (i.e., 1.21) was used **and** all other test acceptance criteria were met.

*Data Analysis Revisions*

*Protocol Revision*: If the lowest toxic concentration calculates to be less than 0%, then the bottom values for IC calculations are set at zero (0) for the Hill function analysis.

*Protocol Revision*: If a biphasic toxicity curve was obtained, the $IC_{80}$ and $IC_{50}$ were calculated from the initial toxicity part of the curve (the $IC_{20}$ was not determined).

*Protocol Revision*: The requirement for test articles to fit the Hill equation with $R^2 \geq 0.90$ was rescinded. The Hill equation was used to characterize the reference substance response curve shape rather than establish acceptance criterion. The PC acceptance criterion was modified to $R^2 \geq 0.85$.

## 2.7 Differences in 3T3 and NHK NRU Test Method Protocols and the *Guidance Document* Standard Protocols

### 2.7.1 Optimization of the *Guidance Document* Protocols Prior to Initiation of the Study

As the NICEATM/ECVAM validation study progressed through Phases I and II, the protocols provided in the *Guidance Document* (ICCVAM 2001b) were optimized to address problems that were encountered. Changes to the *Guidance Document* protocols are described below.

- 3T3 cell seeding density for 96-well plates was increased from $1x10^4$ cells/well to $2.0 – 3.0x10^4$ cells/well to achieve adequate cell growth.
- The calcium concentration in NHK medium was changed from 0.15 mM to 0.10 mM. The test laboratories had expressed concern that cell differentiation would occur at the higher concentration and requested a lower concentration. CAMBREX Clonetics®, the supplier of the NHK cells and NHK medium used in this study, normally grows NHK cells in 0.15 mM calcium without differentiation issues. The supplier agreed that the cells would grow well at 0.10 mM but should not be cultured at concentrations < 0.10 mM in order to avoid morphology and growth changes (CAMBREX technical division, personal communication).
- NHK cells were subcultured once (rather than the three passages suggested in the *Guidance Document*). The laboratories expressed concern about differentiation occurring in the cells if kept in culture too long.
- The highest final concentrations of DMSO and ETOH in the culture media were reduced from 1% to 0.5%. IIVS performed experiments with both cell types to determine the appropriate solvent concentration to avoid toxicity. 3T3 cells

were tested with ETOH at 0.5, 1, and 2% concentrations and DMSO at 0.1, 0.2, 0.3, 0.4, 0.5, 1, and 2% concentrations. The 0.5% concentrations of both solvents were chosen as optimal since that concentration of ETOH produced no toxicity. Although 0.5% DMSO produced slight toxicity (i.e., cells were 91% viable as compared to the control cells – See **Appendix E**), it was chosen by the SMT and laboratories as an acceptable trade off between slight toxicity and the ability to reference substances at higher doses and was used throughout the study (see Curren et al. 2003). However, ETOH was not used as a solvent in the NICEATM/ECVAM validation study.

- The pH of reference substance solutions was not adjusted with NaOH or HCl regardless if solutions became acidic or basic (optimum mammalian cell culture pH is ~ 7.4 [Freshney, 2000]) since some of the basal cytotoxicity produced by these reference substances may be due to pH extremes. See **Appendix F** for pH values of reference substances in culture medium.
- The $CO_2$ concentration in the incubator was reduced from 7.5% *(Guidance Document*) to 5.0% since the laboratories were already set up to use 5% $CO_2$ (a typical optimum $CO_2$ concentration for mammalian cell culture).
- Washing and fixing the cells with a formaldehyde solution prior to NR elution from the cells was eliminated. FAL's regulatory waste disposal requirements concerning formaldehyde were an issue and the NR desorb solution (1% glacial acetic acid, 50% ETOH, 49% $H_2O$) adequately fixed the cells to the test plate (INVITTOX 1991). The SMT and laboratories agreed that the use of formaldehyde was unnecessary.
- Reference substance exposure time for 3T3 cells was extended from 24 hours *(Guidance Document*) to 48 hours (see **Section 2.2.6** and **Appendix E**).
- Cell culture seeding densities for subculture were provided as guidelines and the laboratories were given liberty to determine adequate cell densities (see **Table 2-4**).

**Table 2-4 Cell Culture Seeding Densities**

| Protocol | 3T3 cells/cm$^2$ subculture to flasks | 3T3 cells/well 96-well Plate | NHK cells/cm$^2$ subculture to flasks | NHK cells/well 96-well Plate |
|---|---|---|---|---|
| *Guidance Document* | $1.25x10^4$ | $2.5x10^3$ | $3.5x10^3$ | $2 – 2.5x10^3$ |
| Phase Ia | $0.42 – 1.68x10^4$ | $2.5x10^3$ | $2.5 – 9x10^3$ | $2 – 2.5x10^3$ |
| Phase Ib | $0.42 – 1.68x10^4$ | $2.5x10^3$ | $2.5 – 9x10^3$ | $2 – 2.5x10^3$ |
| Phase II | $0.42 – 1.68x10^4$ | $2 – 3x10^3$ | $2.5 – 9x10^3$ | $2 – 2.5x10^3$ |
| Phase III | $0.42 – 1.68x10^4$ | $2 – 3x10^3$ | $2.5 – 9x10^3$ | $2 – 2.5x10^3$ |

### 2.7.2 Optimization of the *Guidance Document* Protocols During the Study

*Changes in Phase Ia*

- To avoid precipitation of serum components, reference substances were dissolved in treatment medium without NCS for the 3T3 NRU test method (*Guidance Document* recommended 10% NCS). The final 5% NCS on cells in the test plate came from the 50:50 dilution of the treatment medium with the 10% NCS in the routine culture medium (see **Section 2.6.1** – *Precipitate Formation*).
- The volume of NHK medium was reduced from 250 µL per well to 125 µL well for cell seeding. Culture medium was not removed prior to reference substance application. Cell death occurred during the refeeding step (see **Section 2.6.1** – *Cell Growth*).
- To avoid NR crystal formation, NR dye concentrations were reduced from 50 µg/mL to 33 µg/mL (3T3) and 25 µg/mL (NHK) (see **Section 2.6.1** – NR Dye Crystals).
- The PC test acceptance criterion for the $IC_{50}$ was changed for 3T3 and NHK cells to historical mean ± 2.5 standard deviations instead of within the recommended 95% confidence interval of historical mean for 3T3 cells and 2 standard deviations for NHK cells.
- The test acceptance criterion for the mean $OD_{540}$ (> 0.3) of the VC was eliminated. The study protocols provided an $OD_{540}$ range as a guideline (see **Table 2-1** and **Section 2.2.9.**).

*Changes in Phase Ib*

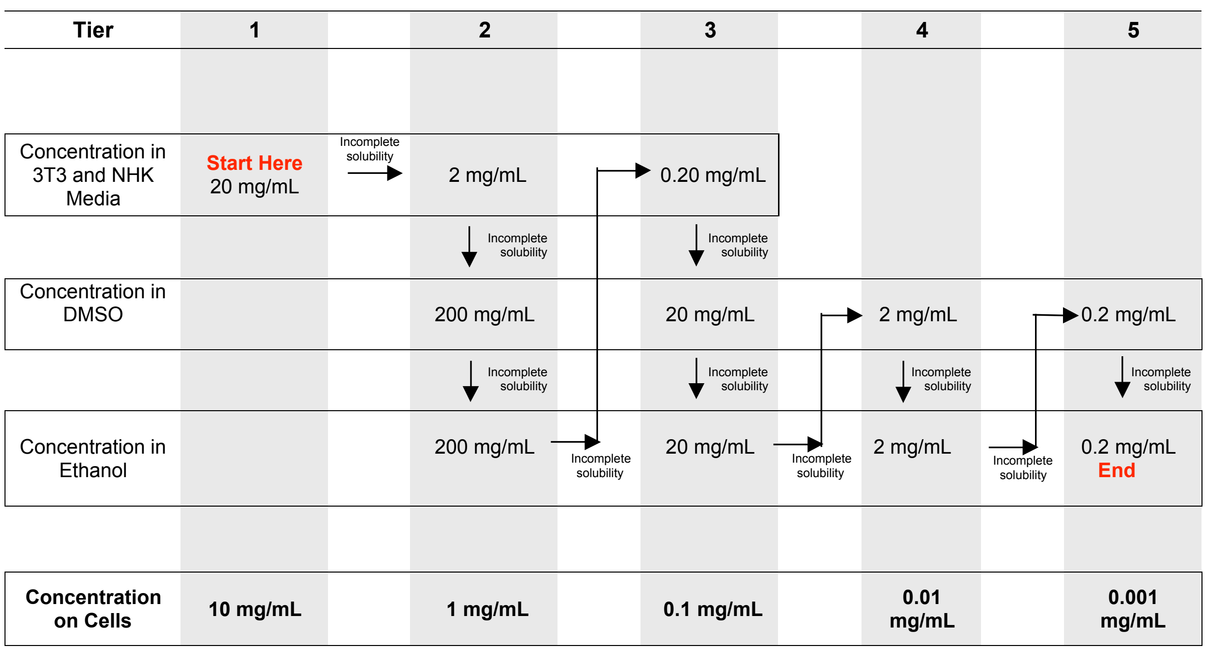
- NHK cells were deemed ready for reference substance application when they reached 20+% confluency rather than the range of 30 – 50% confluency. Laboratory experience in Phase Ia dictated this change.
- A recommendation for obtaining three cytotoxicity points between 10 and 90% inhibition of NRU for use as a quality check of the dose responses was changed to become a test acceptance criterion. The dose response curve had to have at least one calculated cytotoxicity value ≥ 10.0 % and ≤ 50.0 % viability and at least one calculated cytotoxicity value ≥ 50.1 % and ≤ 90.0 % viability (see **Section 2.6.2** – *Test Acceptance Criteria*).
- Instructions for using plate sealers were added to the protocols for testing volatile reference substances (see **Section 2.6.3** – *Testing Volatile Reference Substances*).

## 2.8 Overview of the Solubility Protocol

The SMT, with assistance from the laboratories, developed a solubility protocol to provide information to the laboratories to optimize the determination of the most appropriate solvent to use among three solvents: culture medium, DMSO, and ETOH. Each laboratory tested the solubility of each reference substance using this protocol and provided the data to the SMT prior to initiating the cytotoxicity testing of each reference substance. The SMT analyzed the solubility data provided by BioReliance and each testing laboratory, then assigned the solvents for each test article for this study. This eliminated potential variability in the NRU test methods that may have been produced if different solvents had been used for testing the same substance between laboratories.

The solubility protocol is based on an EPA guideline (EPA 1998) that involves testing for solubility in a particular solvent, beginning at a relatively high concentration and proceeding to successively lower concentrations by adding more solvent as necessary for dissolution. Testing stops when, upon visual observation, the procedure produces a clear solution with no cloudiness or precipitate. The solubility protocol used by the *in vitro* laboratories during Phase III required testing reference substances in the various solvents at equivalent reference substance concentrations applied to the cultures. The solubility flow chart in **Figure 2-7** shows, for example, that 2 mg/mL medium and 200 mg/mL DMSO or ETOH were equivalent concentrations since they yielded 1 mg/mL in cell culture. When applied to cultures, medium was diluted by one-half. The 0.5% [v/v] final concentrations were achieved by diluting DMSO and ETOH by 200. At each concentration, the following mixing procedures were employed, as necessary, to completely dissolve the reference substance in this order: vortex (1–2 minutes); sonication (up to 5 minutes); warming to 37°C (5 – 60 minutes [NRU protocols allow warming to be extended to three hours if cytotoxicity in the range finder test was limited by solubility]). If the reference substance was still undissolved, the next concentration/solvent was tested.

**Figure 2-7 Flow Chart for Determination of Reference Substance Solubility in Medium, DMSO, or ETOH**




Notes: 3T3 Medium - DMEM (Dulbecco's Modification of Eagle's Medium) with supplements; NHK medium - KBM® (Keratinocyte Basal Medium) with supplements (from CAMBREX Clonetics®).

## 2.9 Components of the Solubility Protocol

### 2.9.1 <u>Medium, Supplies, and Equipment Required</u>

*Medium and Chemical Supplies*

- <u>3T3 Cell Medium</u>: DMEM without L-Glutamine and containing Hanks' salts and high glucose [4.5gm/l]; L-Glutamine 200 mM; NCS
- <u>NHK Cell Medium</u>: Keratinocyte Basal Medium without $Ca^{++}$ (KBM®, Clonetics® CC-3104); KBM® SingleQuots® medium supplements (Clonetics® CC-4131) -- epidermal growth factor, insulin, hydrocortisone, antimicrobial agents, bovine pituitary extract; Calcium SingleQuots® (Clonetics® CC-4202); penicillin/streptomycin solution
- U.S.P. analytical grade DMSO
- U.S.P. analytical grade (100%, non-denatured) ETOH

*Equipment*

- waterbath (37°C)
- sonication unit
- vortex unit
- pippettors (micropipettors)
- balance
- pH meter

*Procedures*

The first Phase III solubility protocol procedure was the dissolving of ~10 mg of a reference substance in ~ 0.5 mL medium (both 3T3 and NHK media were tested) for a concentration of 20 mg/mL (see **Appendices B-1 and B-2**). In order, the mixture was vortexed for 1-2 minutes, sonicated for up to 5 minutes, and warmed to 37°C for 5-60 minutes as necessary to dissolve the reference substance. The endpoint for dissolution was that a clear, not cloudy solution with no noticeable precipitate. If the reference substance was not soluble in medium at 20 mg/mL, then more medium was added to a concentration of 2 mg/mL (i.e., a total volume of ~5 mL) (Step 2). The mixing procedures were repeated as necessary to dissolve the reference substance. If the reference substance was not dissolved, ~10 mg reference substance in ~0.5 mL DMSO was added in an attempt to dissolve it at 200 mg/mL DMSO (Step 3). If the reference substance was not dissolved, the same concentration was attempted in ETOH (Step 4). Step 5 began in the same way with 0.2 mg/mL medium and then to 20 mg/mL DMSO and then 20 mg/mL ETOH.

Determination of solubility of reference substances was limited to visual observation of the reference substance in solution. If a solution appeared clear, then solubility testing ceased. If particles were visible or the solution appeared cloudy, then more stringent mixing procedures were employed. If necessary, the solubility procedure proceeded to the next solvent/concentration tier. The duration of the solubility test was dependent on mechanical procedures used to achieve solubility. Some reference substances were immediately solubilized (e.g., liquids) and others required up to 60 minutes of heating and other mechanical procedures.

### 2.9.2 <u>Data Collection</u>

All laboratories (including the reference substance distribution laboratory [BioReliance]) used a worksheet designed to capture the solubility information for the reference substances. The protocol's tiered approach to determining solubility of each reference substance was followed. The endpoint for each step was a visual observation of the solution and a documented comment of soluble or insoluble. Each worksheet contained:

- reference substance code and physical description
- solvent (3T3 medium, NHK medium, DMSO, ETOH)
- amount of reference substance (mg)
- volume of solvent added and total volume (mL)
- concentration (μg/mL)
- pH and solvent color
- mechanical procedures (vortexing, sonication, heating)
- comments (soluble/insoluble at the particular concentration; visual observations)

The solubility test data from the laboratories were transferred via email to the SMT and stored on the NICEATM server and as hard-copy printouts. Each laboratory also maintained electronic and hard-copy files of the data.

### 2.9.3 Variability in Solubility Measurement

Solubility analyses were not replicated since within-laboratory results were not expected to vary. Comparison of the laboratory results to determine laboratory concordance for the 72 reference substances (see **Section 4** for results) provided a measure of variability among the laboratories (see **Section 7**).

### 2.9.4 Solubility and the 3T3 and NHK NRU Test Methods

Reference substance solutions were monitored throughout all aspects of the *in vitro* NRU cytotoxicity test methods and observations were documented. The 2X and 1X solutions for the range finder tests were permitted to contain precipitates. The lowest concentration of reference substance in a 2X solution that contained observable precipitates, particles, globules, or oily droplets was noted in the EXCEL® template. After reference substance exposure, all wells of the 96-well test plates were observed microscopically and scored using a visual observation code as per the NRU protocol. The code addressed growth characteristics and the presence or absence of precipitates. The Study Directors made determinations of test acceptance based on the effect that precipitates had on the NRU results.

### 2.9.5 Methods for Analyzing Solubility Data

During Phase III, the SMT used the solubility data from all the laboratories to determine the solvent that would be used for cytotoxicity testing (see **Section 5** for solubility results and SMT selections). If the solubility of an individual reference substance in 3T3 medium and NHK medium was different, the SMT chose the same solvent for both test methods, rather than choosing one for the 3T3 NRU test method and one for the NHK NRU test method. For example, if solubility in one medium was ≥ 2 mg/mL and solubility in the other medium was < 2 mg/mL, and the reference substance was soluble in DMSO at 200 mg/mL, then the SMT selected DMSO as the solvent for cytotoxicity testing. Where possible, the SMT chose a solvent such that all cytotoxicity laboratories could obtain solubility at some concentration. For example, if a reference substance had low solubility in medium (i.e., 2 mg/mL) at one laboratory and high solubility in DMSO at the other laboratories, the SMT chose DMSO.

Solubilizing enough reference substance to produce cytotoxicity was challenging for relatively insoluble low toxicity reference substances such as lithium carbonate (in the 3T3 NRU test method) but generally was not a problem for toxic reference substances. Some insoluble and highly toxic reference substances were problematic, however, because the amount of powdered reference substance added to solvent was very small, so it was difficult to determine the absence of solute particles in solution (i.e., if the solution was visibly clear). Any undissolved reference substance remaining might have been too little to see. Arsenic trioxide is an example of such a solute.

## 2.10 Basis of the Solubility Protocol

The solubility protocol used by BioReliance, which tested solubility of the reference substances prior to testing by the *in vitro* laboratories, is provided in **Appendix G**. The protocol is based largely on information from the literature and Internet searches for solubility procedures, the experience of the SMT and IIVS, and the solubility and $IC_{50}$ information for the RC chemicals (Halle 2003). The only formal solubility protocol discovered was the EPA Product Properties Test Guideline, OPPTS 830.7840 Water Solubility Column Elution Method; Shake Flask Method (EPA 1998).

### 2.10.1 <u>Initial Solubility Protocol Development</u>

BioReliance tested reference substances in cell culture media at 2000 mg/mL, 400 mg/mL, and 200 mg/mL, and if not soluble, in DMSO, and then ETOH at the same concentrations (initial protocol). It was apparent that these concentrations were not low enough when the laboratory was unable to achieve solubility for arsenic trioxide. The solubility protocol was revised twice to lower the concentrations tested (see **Table 2-5**). An extra tier of concentrations ≤ 1 mg/mL was added for insoluble reference substances. Because of this experience, this solubility protocol for the cytotoxicity laboratories was revised to reduce the number of steps required (by testing in log units) and to test in tiers in which the reference substance concentrations reflected the same concentrations in cell cultures.

In Phases Ib and II, the SMT used the data from BioReliance to determine the solvent for the *in vitro* laboratories to use for NRU testing. When it became apparent that the laboratories sometimes obtained different results than those reported by BioReliance, the SMT used the cytotoxicity results from all the laboratories to determine the solvents for Phase III reference substances.

**Table 2-5 Comparison of Concentrations Tested in Various Solubility Protocols**

| Solubility Protocol Version | Concentrations Tested (mg/mL) | | | | | |
|---|---|---|---|---|---|---|
| | Step 1 | Step 2 | Step 3 | Step 4 | Step 5 | Steps 6-10 |
| BioReliance 1 (4/26/02) and Phase Ia for cytotoxicity laboratories | 2,000 | 400 | 200 | | | |
| BioReliance 2 (9/17/02) | 200 | 40 | 20 | 10 | 2 | |
| BioReliance 3 (10/11/02) | 200 | 40 | 20 | 10 | 2 | 1, 0.5, 0.25, 0.125, 0.05 |
| Phases Ib, II, III for cytotoxicity laboratories | 20 Medium | 2 Medium 200 DMSO 200 ETOH | 0.2 Medium 20 DMSO 20 ETOH | 2 DMSO 2 ETOH | 0.2 DMSO 0.2 ETOH | |

DMSO – dimethyl sulfoxide
ETOH – ethanol
Medium – cell culture medium

The protocol provided a tiered approach for determining the 2X stock concentration for each reference substance, based on the solvent and solubility of the reference substance (see **Figure 2-7**). The solubility protocol was developed to reduce the number of steps for testing (compared to that used by BioReliance) so that solubility testing was less time consuming (see **Appendix B-3**).

### 2.10.2 Basis for Modification of the Phase II Protocol

All three cytotoxicity laboratories found arsenic trioxide (tested in Phase Ib) less soluble than that reported by BioReliance (0.25 mg/mL in 3T3 medium and 0.05 mg/mL in NHK medium). Use of the solubility procedures in the protocol did not dissolve arsenic trioxide. IIVS warmed the stock solution (at least 200 μg/mL for 2X) for longer than the protocol specified (i.e., 30-50 min) but still had small, undissolved particles persist in the non-homogeneous stocks (i.e., particles readily fell out of suspension). ECBC obtained a clear solution (highest 2X concentration was 30-50 μg/mL), but found precipitated particles after the solution stood at room temperature. Sonication time was increased to 15-30 min, and heating time to ~ 30 min to get a finer suspension. This procedure achieved a more homogeneous mixture, resulting in better series dilutions and more uniform application of reference substance to the cells. FAL stirred the suspension (~ 20-90 μg/mL) in the $CO_2$ incubator for 1.5 to 2 hours to get clear medium.

*Protocol Revision for Phase II*: The duration of the solution heating range was increased from 5-20 minutes to 5-60 minutes.

### 2.11 Summary

- The *Guidance Document* NRU protocols were the basis of the NICEATM/ECVAM study protocols. The SMT and laboratories made initial modifications to the protocols prior to implementation of the study. Other protocol modifications were made after commencement of testing and were the result of comments and recommendations from the laboratories and the SMT. The resulting optimized protocols were used in the main testing phase (Phase III) and were the final protocols for the NICEATM/ECVAM study.
- The solubility protocol was developed to provide specific guidance to laboratories to assure that solubility issues could be satisfactorily addressed and reference substances from a specific study set could be adequately prepared and evaluated for *in vitro* cytotoxicity effects.

***[This Page Intentionally Left Blank]***